REDES NEURAIS COM TENSORFLOW

DIEGO RODRIGUES DSC

INFNET

Agenda

Parte 1 : Meta Heurística de Treinamento Robusta II

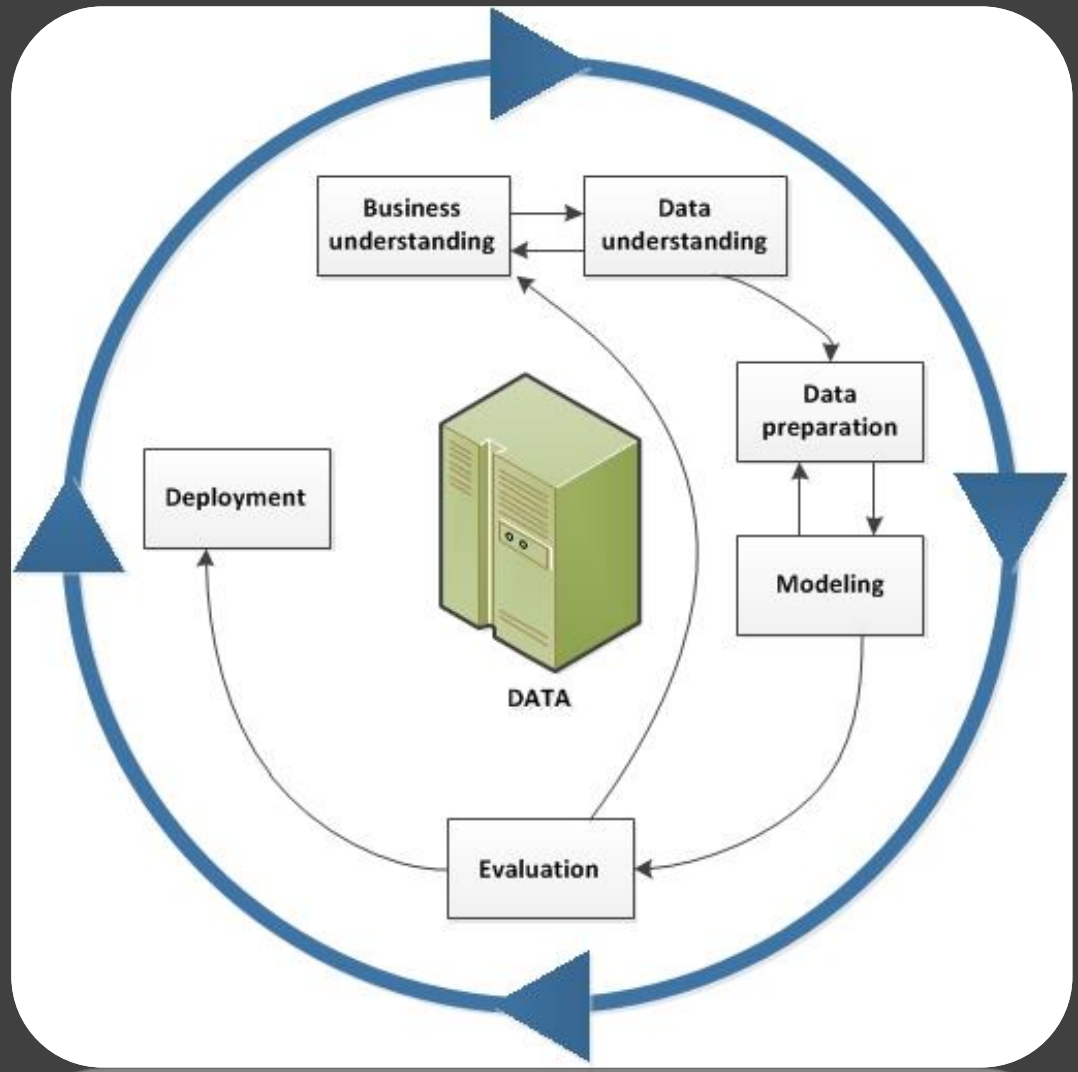
- Novo ciclo do CRISP
- Regressão / Aproximação
- Tensorflow Time Series Tutorial
- Validação & Figuras de mérito para Regressão

Parte 2 : Prática

- Notebook: Regressão Iris

Parte 3 : Trabalhos

- Escopo & Evolução



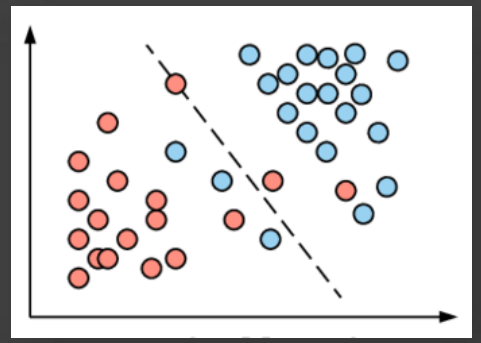
Cross Industry Process for Data Mining (CRISP-DM)

Novo Ciclo CRISP

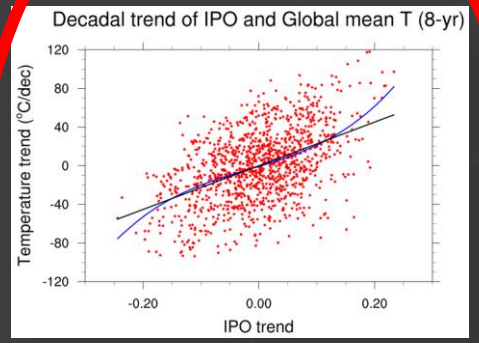
Algoritmo	Representação	Preparação	Modelagem	Validação
• Reta 2 Pontos • NN 10% VAL • NN 10 Folds • PS10 • PS10	• 2D • 2D • 2D • 4D / 3 Classes • 7D / 1D Regressão	• Nenhuma • Nenhuma • Scale • Scale • Scale	• Reta 2 Pontos • 1 Neurônio • 1 Hidden • 1 Hidden • 1 Hidden	• Nenhuma • Precisão/Recall • Precisão/Recall • Acurácia • MSE

- Regressão multivariada com uma variável categórica
- Busca nos hiperparâmetros ótimos (# funções de ativação)
- Identificar as variáveis mais relevantes

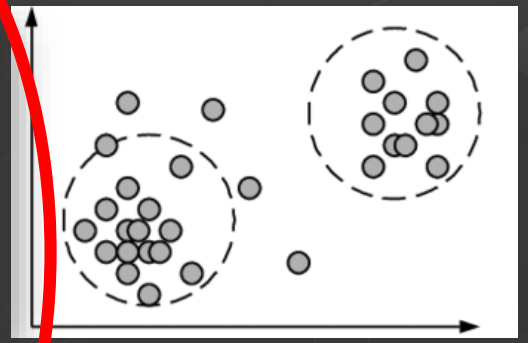
Paradigmas de Modelagem Estatística



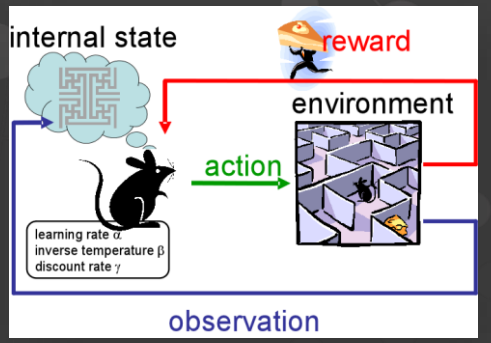
SUPERVISIONADO – CLASSIFICAÇÃO



SUPERVISIONADO – REGRESSÃO

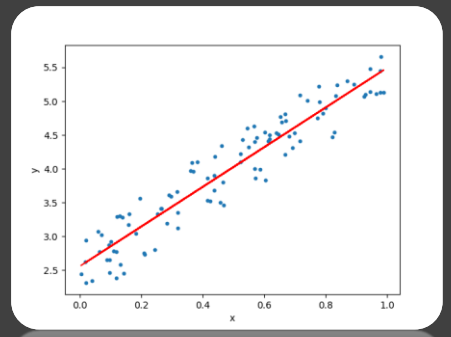


NÃO SUPERVISIONADO

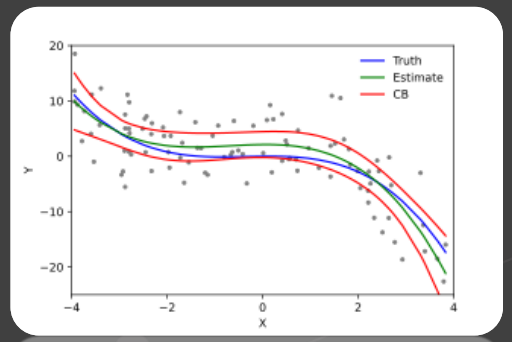


APRENDIZADO POR REFORÇO

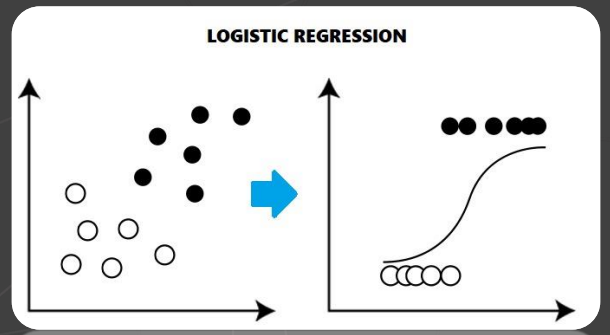
O aproximador Universal $Y = F(X) + \varepsilon$



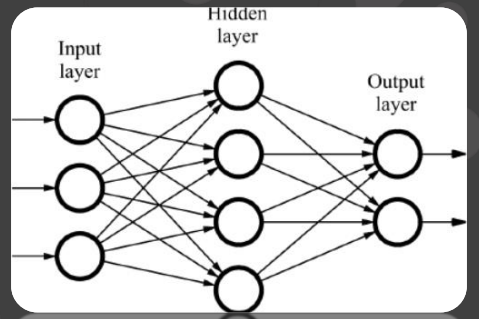
$$Y = \alpha^T x + \varepsilon$$



$$Y = X\alpha + \varepsilon$$

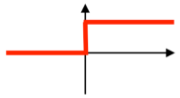
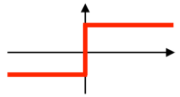
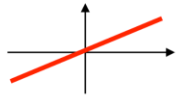
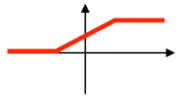
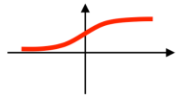
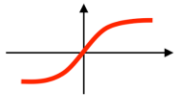
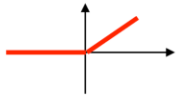
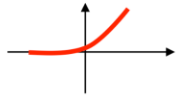


$$Y = \frac{1}{1 + e^{\alpha^t x + \varepsilon}}$$



$$Y = \varphi(x) + \varepsilon$$

Funções de Ativação

Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = max(0, z)$	Multi-layer Neural Networks	
Rectifier, softplus	$\phi(z) = \ln(1 + e^{z})$	Multi-layer Neural Networks	

Tensorflow Time Series Demo

Treinamento

Erro Médio Quadrático

- Minimização do MSE no conjunto de treino, controlado pelo conjunto de validação.
- Estratégia de busca idêntica a de classificação.

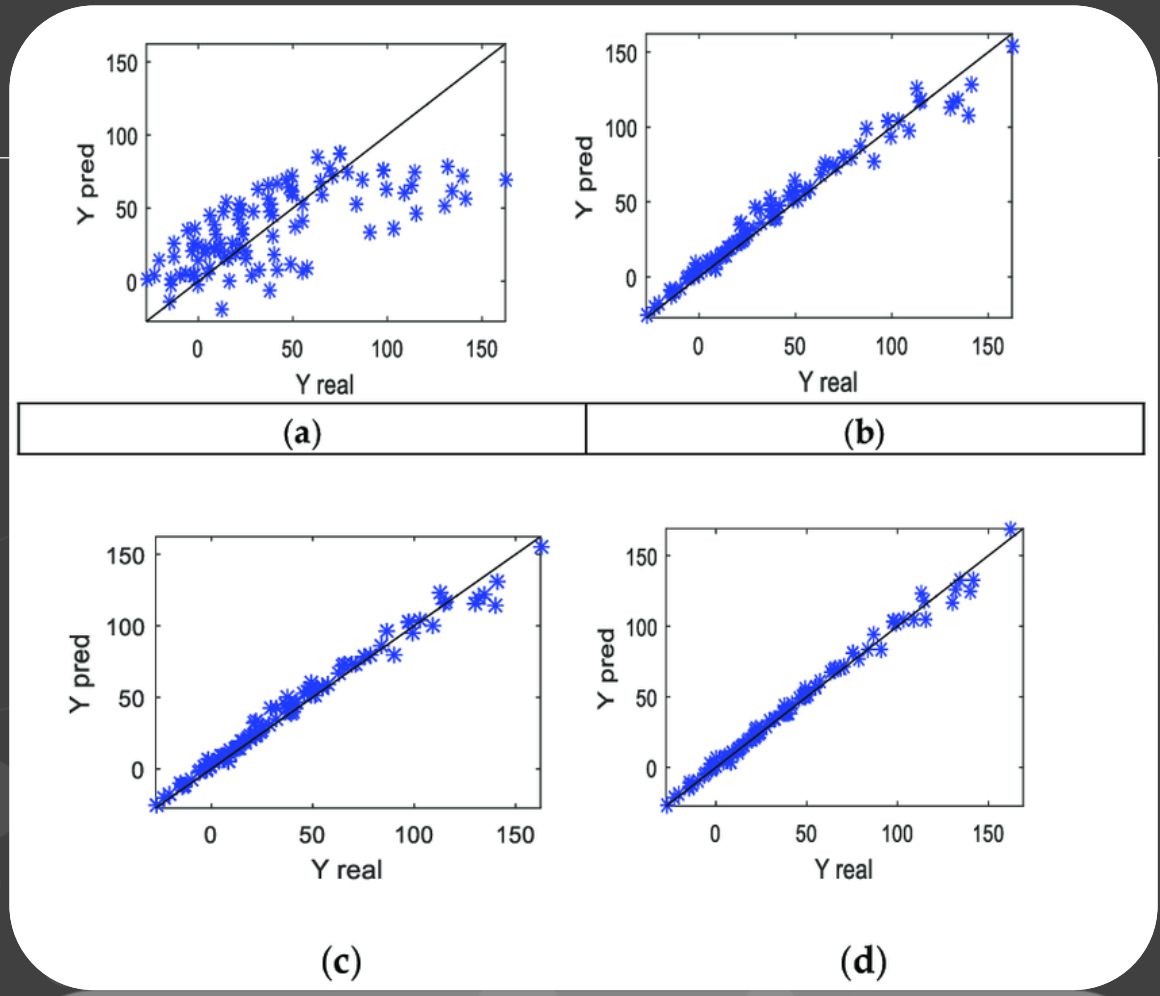
$$MSE = \frac{1}{N}\sum_{i=1}^{N}(f_i - y_i)^2$$

where N is the number of data points,
f_i the value returned by the model and
y_i the actual value for data point i.

Validação

Dispersão Real vs Previsto

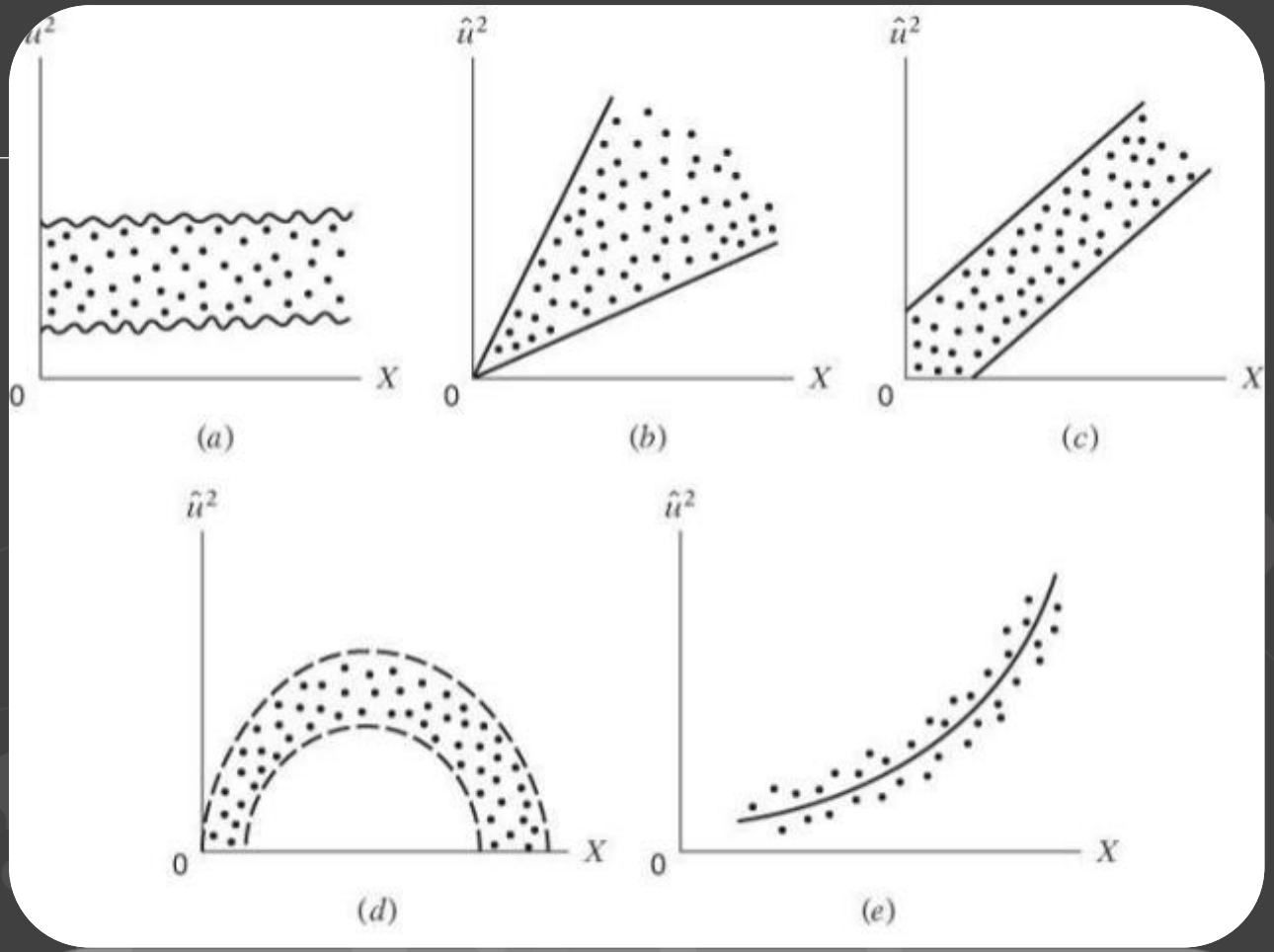
- Queremos os pontos próximo da reta e com distância pra reta constante, independente do Y.



Validação

Dispersão Real vs Previsto

- Modelo Homocedástico / Heterocedástico.



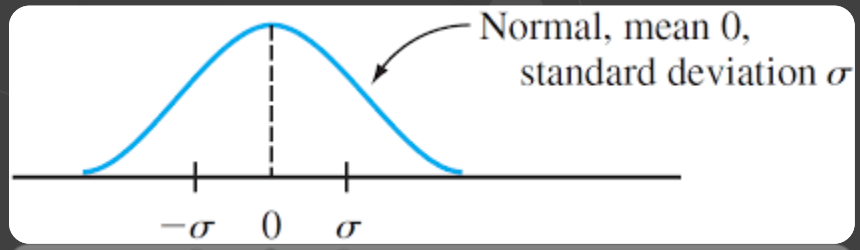
Figuras de Mérito

R Quadrado

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \overline{y})^2}$$

Resíduo Normal de Média Zero e

Variância Constante



Iris Robust Regression

Modelagem

Rede Neural Feed Forward

- Representação: 4 atributos > X atributos mais relevantes
- Hiperparâmetros: PATTERN SEARCH no # de neurônios de cada tipo na camada oculta.
- Treinamento: base de treino completa.
  - MSE
  - Validação Cruzada 10 Folds